DOCUMENT RESUME

ED 285 846 SP 029 262

AUTHOR Van der Sijde, Pieter C.
TITLE The Effect of a Short Teacher Training on Students' Learning Outcomes.
INSTITUTION Twente Univ. of Technology, Enschede (Netherlands).
REPORT NO IST-MEMO-87-02
PUB DATE Apr 87
NOTE 28p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, DC, April 20-24, 1987).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Foreign Countries; Higher Education; Mathematics Teachers; *Program Effectiveness; Secondary Education; Teacher Education Programs; *Teacher Effectiveness
IDENTIFIERS Netherlands

ABSTRACT

A study focused on the effect of a specific mathematics teachers training program on student achievement. Training of 33 eighth-grade mathematics teachers featured instruction in lesson format and management of the classroom and instruction. Observation of the teachers after the training indicated that the time training is given in the course of the school year is an important factor in achieving effectiveness as measured through student achievement. Results also indicated that teachers should be trained, with that training specifically tailored to data obtained from teacher observation. Mailing the training manual appeared not to be effective for self-training. (CB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED285846

TO Toegepaste Onderwijskunde

The effect of a short teacher training on students' learning outcomes

Pieter C. van der Sijde

IST-MEMO-87-02

IST MEMORANDUM

Vakgroep Instructietechnologie

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P.C. van der Sijde

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

UNIVERSITEIT TWENTE

SP 029 262

The effect of a short teacher training on students' learning outcomes

Pieter C. van der Sijde

IST-MEMO-87-02

Paper prepared for presentation at the
Annual Meeting of the American Educational
Research Association, Washington DC,
April 20-24, 1987

Session: International Perspectives on Teachers'
Belief and Evaluation Strategies
(Division K)

Running head: TEACHER TRAINING AND STUDENT ACHIEVEMENT

Abstract

Partly on basis of process-product correlations between teacher behavior and learning outcomes of students a training for math teachers has been developed. In a previous study is established that this training does change teacher behavior.

In this study the focus is on the effect of the training (the training manual and the training session) on the student achievement is investigated. From the results it does not become clear whether or not the training as such is effective, but it appears that the time the training is given in the course of the schoolyear (september/january) is of importance. The training given in january and tailored using observationdata is more effective then a training (untailored) given in september. To mail the training manual to teachers appears not to be effective.

Research on the effects of research-based teachers training has increased recently. Gage (1985) indicates, that at the time of his writing there were 13 investigations of this type completed. This type of research will be refered to as "trainingstudies" in the rest of this paper.

Trainingresearch of this kind is part of the descriptive-correlational-experimental loop design (Rosenshine & Furst, 1973) which means that first procedures are developed to describe the teachingprocess quantitatively (descriptive phase), second the variables to describe the teachingprocess are correlated with the increase of student knowledge and change in attitude towards the particular subject. The result of such a correlational study forms the input for the development of a research-based teacher training. In an experiment this training is tested whether or not it can establish an increase of students' learning outcomes.

This last phase, the experimental research, is of major importance, because in this experiment the instructional (and managerial) design rules, which have been deducted from the correlations, are tested. Of the 13 completed studies (mentioned by Gage, 1985) 12 were successful, which means that in 12 studies the experimental training did cause a change in students' learning outcomes.

The content of such teacher trainings consists out of "teacher-should's", rules to use during classroom management and instruction. Depending on the type of research stress is laid in either of these two types of "teacher-should's". In the investigation of Good ad Grouws (1979; Good, Grouws & Ebmeier, 1983) the accent lies on instruction, while in the investigations of Evertson and Anderson (1979) the accent lies on classroon management.

A problem in designing and constructing research-based trainings is that an important part of the trainig is based on correlational research. The correlations are tranlated into causal relation. Although this problem was already recognized by Fenstermacher in 1976 (Fenstermacher, 1976), it is still important. In Van der Sijde and Tomic (1987) we suggested to study the effects of teachertraining not only in the context of teacher effectiveness (what has been done in all 13 studies mentioned by Gage), but also in the context of teacher competence. In research on teacher competence the effect of a training is studied on the teaching behavior of the participting teachers: the teaching beha vior of teachers is systematically observed before and after the training period, and the dependent variable is the trained teaching behavior.

When a training is developed the first study has to be to investigate whether or not such a training can estblish changes as intended in the actual teaching behavior of teachers. If the result of such an implementation study would be negative, than it is impossible to ascribe increases in students' learning outcomes in an effectiveness study to the training.

In this study we report our research on teacher effectiveness, viz. the influence of the teacher (teaching behavior) on the students' learning otcomes. But before we performed this study we investigated in an implementation study (Van der Sijde, 1987). From this study it appeared that the mean overall increase in trained teacher behavior was about 0.25 standard deviation. Walberg (1984) calls this the effectsize. We concluded from this study that the training we developed is capable of changing teacher behavior and that changes in students' learning outcomes we find in the present study can be asccribed to this teacher training.

METHOD

Subjects

Thirty three teachers volunteered to participate in the experiment. All teachers taught mathematics in the 8th grade using the textbook Sigma. The subjects were assigned to four conditions. To condition 1: n = 13 teachers, condition 2: n = 8 teachers, condition 3: n = 6 teachers and condition 4: n = 6 teachers of which one dropped out. The teachers in the condition 1 and 2 participated in the experiment in the period September till March, while the teachers in the conditions 3 and 4 only participated in the period December till March.

Teacher training

On the basis of the results of the Dutch Classroom Environment Study - correlational phase, on results of experiments and theory a teacher training has been developed (Van der Sijde, 1985, 1987). The content of the training is organized around three topics, the format of the lesson, management of the classroom and instruction. The content of the training is put in a 40-page manual entitled "Management en instructie in wis kundelessen" ("management and instruction in mathematicslessons"). In this manual rules/recommendations for teaching are included (see Table 1).

Table 1

A part of the content of the manual on classroom management is based on Kounin (1971) and Borg (1975) and are recommendations on withitness, learner accountability and group focus. The recommendations on instructions are on explanation, questioning and feedback.

Observers

Nine graduate students (in Education and Science) were trained for about 40 hours to master the observation instrument TOOL-W implemented within the TOOL-system.

Observation instrument: TOOL-W

This is the observation instrument developed to observe the teaching behavior in the classroom. It contains 33 low inference categories. There are 4 codes for the direction of look of the teacher, 6 codes for the lessonphase, 8 codes for classroom management and 15 codes for instructions. Every code is a low inference operationalization of a recommendation. The TOOL-W instrument is implemented within the TOOLsystem (Van der Sijde & Dirksen, 1987 a,b) which is a serie of connecting programs on portable microcomputer to be used for the specification and operational use of observation instruments (for more details see Van der Sijde & Dirksen, 1987).

Training of observers

The training started with an explanation and discussion of the observation categories in the instrument TOOL-W. Further the training consisted out of practical excercises using videotaped and audiotaped situations. The final part of the training was three observations in the classroom situations. During the training on most of the excercises the interobserver agreement was calculated as an indication for the progress of the observers. At the end of the training this interobserver agreement was satisfying (for details, see Van der Sijde, 1986).

Attitude scale

The attitude scale consists out of a selection of items from the student questionaire developed for the IEA's Second Mathematics Study. 18 Items were selected and this new attitude scale has two subscales: enjoying mathematics (containing items such as: "I like mathematics", "I want more mathematics") and self-image (containing items such as: "I'll never be good in mathematics", "I have more trouble with mathematics than others"). Each item had to be rated on a five point scale and for each of the two subscales a score is calculated.

Cognitive test

The cognitive test contains 30 items. All testitems are constructed by the Dutch Central Institute for Test Development (CITO). The pretest was the same for all teachers and their classes. The teachers were given the test to check for opportunity to learn. The post tests differed between teachers, because every test should correspond with the content covered by the teacher. In order to make comparison possible each score is converted into a standardscore by using the mean and standarddeviation of the group.

Procedure

In the experiment there are four conditions:

Condition 1: The teachers in this condition are received at our institute to participate in a one day training course. After the training a cognitive test and an attitude scale administed (pretest). Then 8 to 10 lessons during a 2½ month period are observed by trained observers using the TOOL-W instrument. After the observations the same attitude scale and a cognitive test is administered (first post test) and after another

period of 2½ month again the same attitude scale is administered for the third time together with a third cognitive test (second post test).

Condition 2: This condition resembles condition 1 except for the fact that the teachers do not receive the training before the observation period, but after the observations of the lessons.

Condition 3: The teachers in this condition received the training manual and in their classes only the first and second post tests, each consisting out of the attitude scale and a cognitive test, were administered.

Condition 4: In the classes of the teachers in this condition the first and second post test were administered.

Design and data

The design used in this study can be depicted as:

$$\begin{array}{lll} X\ O_1\ X & & X \\ \text{----------} & & \\ X\ O_2\ X & & X \\ \text{----------} & & \\ & X\ O_3\ X & \\ & \text{------} & \\ & X \quad X & \end{array}$$

In order to study the implementation of the lesson phase model, three variables are constructed. The first is the L-score and L=5- (the number of observed lessons phases); in the ideal lesson L=0. The second variable is E, E=0 when the lesson starts ith phase 1 and otherwise E=1. For every lesson a S-score is calculated which expresses the violations against the lesson phase model. An ideal lesson (according to the lesson

phase model) starts with phase 1, followed by 2,3,4 and 5. Violations against the model are e.g. phase 1 followed by phase 4 (forward violation) or phase 3 followed by phase 1 (backward violation). The ideal lesson has a S-score of S=0. Some violations are valued as minor (from phase 4 to phase 3, or from phase 3 to phase 5), while some are valued as major (from phase 5 to phase 2). S is the third variable to establish the degree of implementation of the lesson phase model.

Further in TOOL-W a number of Kouninian variables are operationalized. In the results only the Kouninian variables are reported.

Withitness is operationalised in low inference variables. These low inference variables are: the teacher atches the pupil (direction of look is toward the pupils), the teacher changes place (he takes an other position in the classroom), the teacher looks at a particular pupil to indicate he has seen the pupil shos inappropriate behavior, the teacher rebukes a pupil and the teacher suggests alternative behavior to a pupil who shows inappropriate behavior. The number of times the teacher performs such behavior is observed before and after the training. Withitness is the standarized sum of the five mentioned low inference behaviors.

Group alerting is operationalized by the number of times the teacher alerts the class, e.g. with "this is important" or "pay attention".

Learner accountability is operationalized by the number of times a teacher asks for progress made by the pupils (e.g. "who has completed ....", "who is not ready by now?").

Results

Interobserver agreement

To establish the agreement between the observers a number of lessons are observed by two observers at the same time. Using Cohens kappa (Cohen,

1960) the interobserver agreement was calculated.

The mean kappa for the direction of look: k= 0.48; for the lesson phases: k= 0.98; for classroom management: k= 0.84; and, for instruction: k= 0.59.

Lesson model

The teacher condition 1 spend significently more time on monitored practice (t_{19} = 1.60, $p < .10$) and on guided practice (t_{19} = 1.30, $p < .10$) - see Table 2.

Table 2

In the training, stress is laid upon the presence of each of the five lesson phases in every lesson. Not in every observed lesson these five phases are present, however there is a significant difference between condition 1 and condition 2 with respect to the L-variable, (X_5^2 = 27.16, $p < .001$) - see Table 3.

Table 3

Further every lesson should start with the review of homework. With respect to the E-variable is no difference between the conditions (X_1^2 = 0.03, p = n.s.). With respect to the S-variable there is a difference between the conditions (t_{189}=-1.49, $p<.10$): A mean score for teachers in condition 1 is 3.12 (standard deviation: 3.40) and for teachers in conditions 2 the score is 4.00 (s.d.: 4.69).

Teacher behavior: management

The mean withitness-score (W) of the teachers in condition 1 is W = 0.26 (sd=2.85) and of the teachers in condition 2 is W = -0.42 (sd=2.00). The difference between the two conditions for withitness is nog significant (MWU= 46, p= ns), the effectsize is E= 0.34. (see note 1).

The mean group alerting-score (GA) of the teachers in condition 1 is GA= -0.05 and of the teachers in condition 2 is GA= 0.31. The difference is not significant (MWU= 71, p= ns), the effectsize is E= .37.

The mean score for learner accountability (LA) of the teachers in condition 1 is LA= 0.30 (sd= 1.08) and of the teachers in condition 2 is LA= -.037 (0.73). The difference between the conditions is significant (MWU= 26.5, p< .10).

Teacher behavior: instruction

Instruction is divided into three components, explanation, question and feedback. According to the recommendation in the teacher training the phase "presentation" should start with a review of relevant concepts. The teachers in condition 1 spend on the average 0.55 seconds (sd= .50) on a review per lesson, while the teachers in condition 2 spend 0.25 seconds (sd=.30) on a review, the effectsize is E= 1.00. Instead, it is also possible to have students answer questions on prerequisites, the teachers in condition 1 spend on the average 2.75 seconds (sd= 1.95) while the teachers in condition 2 spend 2.55 seconds (sd= 1.6), the effectsize is E= 0.13. Another aspect of explanation is the use of the textbook, the teachers in condition 1 refer on the average 0.55 seconds (sd= .45) to the textbook and the teachers in condition 2 do this 0.30 seconds (sd= 0.4), the effectsize is E= 0.63. Instead of refering to the textbook the teacher can also indicate (e.g. on the blackboard) what the

important issues are. The teachers in condition 1 do this on the average 0.65 seconds (sd= 0.3) per lesson and the teachers in condition 2 0.60 seconds (sd= 0.55) (Effectsize E= 0.09). The use of examples is stressed in the training and the teachers in condition 1 spend 0.25 seconds (sd= 0.25) and the teachers in condition 2 0.20 seconds (sd= 0.25) - E= 0.2. Presentation with a summary. This summary is conducted by the teachers in condition 1 for 0.30 seconds (sd= 0.35) and by the teachers in condition 2 for 0.60 seconds (sd= 0.70) - E= -0.43.

Questions form a second component of explanation. In the training a distinction is made between lower-order and higher-order questions (accor ding to the taxonomy of Bloom, lower-order questions are questions on knowledge and concepts, while the higher-order questions are on application, analysis, synthesis and evaluation. The teachers in condition 1 are engaged for 10.3 seconds (sd= 4.3) and the teachers in condition 2 for 9.65 seconds (sd= 4.2) in asking lower-order questions (E= 0.15). For asking higher-order questions, the teachers in condition 1: 0.55 seconds (sd= 1) and the teachers in condition 2: 0.50 seconds (sd= 1.1) - E= 0.05.

The third component is feedback. The teachers in condition 1 spend more time on giving positive feedback, E= 0.63 and less time on giving negative feedback, E= 0.21.

Effects of the training on students' learningoutcomes

In the experiment there are four different groups. From the results on the pretests it appears that the results of the teachers in condition 1 are comparable with the results of the teachers in condition 2 (attitude scores for pleasure: Mann-Whitney U = 50; attitude scores for selfimage: MWU = 50; scores on the cognitive test: MWU = 47; none of these Mann-Whitney U-scores is significant). Comparing the results of the first

posttests for the teachers in the two experimental conditions with those of the teachers in conditions 3 and 4 yield no significant difference (attitude scores for enjoying: $X_2^2 = 0.65$; attitude score for selfimage: $X_2^2 = 3.36$; scores on the cognitive test: $X_2^2 = 0.45$. NB. The differences are assessed using the Kruskal-Wallis test). Neither are there differences between the teachers in condition 3 and condition 4 (attitude scores for enjoying: MWU = 13; attitude sores for selfimage: MWU = 16; scores on the cognitive test: MWU = 13). The differences in the results on the second posttest yields no significance (attitude scores for enjoying: $X_3^2 = 1.48$; attitude scores for selfimage: $X_3^2 = 1.49$; scores on the cognitive test: $X_3^2 = 2.81$. The differences are assessed using the Kruskal-Wallis test).

To compare the increase in score from the pretest to the first posttest both scores are standardized on the mean of the group. The differences in increase of the score are not significant different for the attitude scores (enjoying: MWU = 39.5; selfimage: MWU = 46.5) nor for the score on the cognitive test (MWU = 40).

The increase in score from the first to the second posttest does not significantly differ for the teachers in the treatment conditions (condition 1 and 2) and the teachers in the two other conditions (attitude scores for enjoying: MWU = 109.5; attitude scores for selfimage: MWU = 89; scores on the cognitive test: MWU = 111). Neither is there any increase in scores between the four conditions (attitude scores for enjoying: $X_3^2 = 1.15$; attitude scores for selfimage: $X_3^2 = 2.23$; scores on the cognitive test: $X_3^2 = 0.28$ - Kruskal-Wallis test).

When the increase in scores on the attitude scales from the pretest to the second posttest are compared, there is no significant difference (pleasure: MWU = 33; selfimage: MWU = 40), on the cognitive test, however, there is a significant difference in increase (MWU = 27, $p < .10$) in favour of condition 2.

Discussion

It appears from the results of the interobserver agreement, that there is a reasonable agreement between the observers. This means that the observations are reliable and it is allowed to draw conclusions from the data sampled in this experiment. The results show that the mean results of the students of the teachers in condition 1 and 2 are comparable for attitude as well as for the achievements-scores, they do not significantly differ. This holds for the pretest as well as for the first and second post test. This is also the case for conditions 1, 2, 3 and 4 on the two post tests. From these results it can be deduced that the observation of lessons by observers dot not affect the classroom results (no Hawthorne effect).

To access the effectiveness of the training are not the results on the posttests of importance but the increase in score from pretest to posttest(s). Comparing the difference in increase in score for the teachers in condition 1 with those in condition 2, there is no significant difference. This is a remarkable result because such a result has not been reported sofar in other trainingstudies. This result may arise from a difference in design. Most trainingstudies use a pretest-posttest control group design, in such a design the control group is a control group which means that this group will not receive the experimental treatment (the training). In this study there was no "real" control

group. For example, it was clear to the teachers in condition 2 that they would receive the training after the observation period and that those observations would be used to tailer the training. This changes the "control" group into another experimental group. It also changed the intention of the observation: for the teachers in group 1 it was to establish whether or not they implemented the training and for the teachers in group 2 were observed to sample information on their teaching behavior. For this reason it was of importance that an implementation study (a study on "teacher competence") preceded this study (Van der Sijde, 1987). From this study it appeared that the training influences teacher behavior and in this study there are hardly any significant differences in teacher behavior, although some behaviors have a rather great effectsize. If this implementation study was not performed, it would not be possible to establish a realistic degree of implementation on basis of this study: it would be overestimated! It shows once again that an implementation study should be incorporated in the descriptive-correlational-experimental loop design. A conclusion from this study on the implementation of the training would be, that parts of of the training are implemented and parts are not implemented and the overall implementation is satisfying. This conclusion is further supported by the fact that there is an effect of the training on students' achievement. Because of an implementation study has been performed, a more sophisticated conclusion can be reached: First, the implementation studied showed that the training could be implemented, which means that there is no need investigate implementation in this study. Second, a proper design for studying implementation is a one-group pretest-post-test design. In this study the pre-training behavior was not established (just as in almost all training studies), which means that from a study

like this no conclusion with reference to implementation can and should be made. Nevertheless it is important to establish the differences in the teaching behavior between the two groups (teachers in condition 1 and condition 2), but such a comparison should not be confused with implementation. What can be concluded from such a comparison is that some teaching behaviors are so specific for the training that only the teachers in condition 1 could implement them, the teachers in condition 2 practiced such behavior more or less (not all teachers ofcourse) randomly (no implementation).

The teachers in condition 1 implement the lesson phase model, which appears from a lower S- and L-variable, there is no significant difference for the E-variable, which means that most teachers (irrespective of the experimental conditions) start their lessons with a check of homework. The teachers in condition 1 also spend time on monitored practice (from table 3-2 can be deduced that this is approximately 9 minutes per lesson), while the teachers in condition 2 significantly spend less time on this phase (± 5 minutes per lesson). The same holds for guided practice (condition 1: ± 3.5 minutes, and condition 2: ± 2 minutes). It appeared from the conversations with the participating teachers that in practice the discrimation, although in theory rather clear, is very hard to make. The difference to them is of a rather gradual nature. According to their on subjective observation they spend more time on guided practice than on monitored practice, the data shows a different picture. On the other hand when the two lesson phases are considered as one, then the teachers in condition 1 spend ± 11.5 minutes per lesson on the phase and the teachers in condition 2 only ± 7 minutes, while the lesson phase model advises ± 15 minutes. It can be concluded that there are differences in teaching behavior between the teachers in group 1 and group 2.

Because the pretraining behavior of the teachers in group 1 is unknown and the results of many studies indicate that teachers make a difference (e.g. Good, Biddle & Brophy, 1975), it would be inappropriate to deduce from these results a measure to determine the degree of implementation. The difference in the increase in cognitive scores from the pretest to the second post test is significant and in favour of condition 2. This is again a remarkable result. In the first place, because the effect of the training is only noticable on the increase in scores on the cognitive test (student achievement) and second this is only the fact in the period between the pretest and the second post test. This means that in this experiment the question to the effectiveness of the training as such remains unanswered and only an answer to the question on the most effective (training) condition can be given. The most effective condition is the one with the training after the observation period. There is an explanation possible for this effect. The teachers in condition 1 participated in the training, while nothing about their actual teaching behavior was known to the trainer (the author). Many of the recommendations and rules are very logically and in a sense familiar to the teacher, because some of them express what a good teacher <u>should</u> do (and all participating teachers are good teachers, so ...). The teachers in condition 1 were allowed to mix up what they <u>do</u> in the classroom and what they <u>should do</u> in the classroom. Because of this confusion no indications for changes could be given (no data of classroom observation were available). The teachers in condition 2, however, were not allowed to confuse their actual teaching behavior with a desired teacher behavior. During the training all observation data were available and used: the teachers were confronted with the results of the classroom observations. As a result concrete indications for change could be given. To

have the best of both sides it will be probably even more effective when a training is given on two occasions: a first training before a period of observations and a second training after a number of observation of lessons in which the same training, but tailored, is given to the teachers after 2 or 3 months. A number of training studies had a second training session, but not in the sense as meant above. In the training studies by Anderson et al (1979) and Good and Grouws (1979) the period of time between trainingsession 1 and 2 was only 1 week.

When the teachers in condition 3 and 4 are compared on the increase in score from posttest 1 to posttest 2, the results obtained by Coladarci and Gage (1984) are replicated. Using, what they call a minimum intervention, which mean that one of the two groups of teachers is sent a training manual by post. In the study performed by Coladarci and Gage (1984) no significant differences were found. In this study there were no differences found either. There is a difference, however. In the Coladarci and Gage study the teachers were asked to participate in the study. Upon their agreement they received by post the manual(s). In this study the teachers in condition 3 received the manual and were asked to read and practice it. There was no control what so ever. Even when a not-so minimum intervention is added, using a design with 3 conditions, one with a training (condition 2), one with a training manual (condition 3) and a control condition (condition 4), there is no significant difference.

Probably the minimum intervention in this study has been too minimum with respect to the degree of participation and to time. A period of 2½ month (also including holidays) is too short to notice any influence on the students' learningoutcomes. A period of 2½ month, however, is not too short to notice changes in teaching behavior (Van der Sijde, 1987).

Under influence of the training a teacher changes his teaching behavior and changes do not take place from one day to the other but it takes time.

It is once more established that even short (1 day) training course can succesfully change teacher behavior and through this intervention student achievement.

Notes

Note 1: Effectsize is defined as $E = (X_A - X_B)/ sd_B$, where X_A is the mean score condition 1, X_B is the mean score condition 2 and sd_B is the standard deviation of X_B (Walberg, 1984).

References

Borg, W. (1975). Instructional handbook for the protocol modules on classroom management. Logan: Utah University.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Coladarci, T., & Gage, N.L. (1984). Effects of a minimal intervention on teacher behavior and student achievement. American Educational Research Journal, 21, 539-555.

Gage, N.L. (1985a). Hard gaines in the soft sciences. Bloomington: Phi Delta Kappa.

Gage, N.L. (1985b). What educators should know about research on teaching. Paper prepared for the Doctor Residency Seminar, College of Education, University of South California.

Gage, N.L., & Giaconia, R. (1981). Teaching practice and student achievement: causal connection. New York University Education Quarterly, 12, 2-9.

Good, T.L., Biddle, B.J., & Brophy, J.E. (1975). Teachers make a difference. New York: Holt, Rinehart & Winston.

Good, T.L., & Grouws, D.A. (1979). The Missouri mathematics effectiveness project. An experimental study in fourth grade classrooms. Journal of Educational Psychology, 71, 355-362.

Joyce, B., & Weil, M. (1972). Models of teaching. London: Prentice Hall.

Kounin, J.S. (1970). Discipline and group management in classrooms. New York: Holt, Rinehart & Winston.

Rosenshine, B., & Furst, N. (1973). The use of direct observation to study teaching. In R.M.W. Travers (ed), Second handbook of research of teaching. Chicago: Rand McNally.

Van der Sijde, P.C. (1985). The implementation of a research-based teacher training. Enschede: University of Twente.

Van der Sijde, P.C. (1986). Training van observatoren t.b.v. het hoofdonderzoek (Training of observers for the main study). Enschede: University of Twente.

Van der Sijde, P.C. (1987). Development and implementation of a research-based teacher training. Submitted for publication.

Van der Sijde, P.C., & Dirksen, A. (1987). The TOOL-system. In Maarsse, F.J. et al (eds), Computers in Psychology, methods, instrumentation and psychodiagnostics. Lisse: Swets & Zeitlinger.

Walberg, H.J. (1984). Improving the productivity of America's schools. Educational Leadership, 41, 19-27.

Table 1

A summary of instructional and managerial rules

Lesson phase model:	phase 1	Review of homework	(5 à 10 min)
	phase 2	Presentation	(10 à 15 min)
	phase 3	Monitored practice	(± 5 min)
	phase 4	Guided practice	(± 10 min)
	phase 5	Homework/tutoring	(± 10 min)

Classroom management

Use a beginning-of-the lesson signal

Make transitions between two lesson phases as smooth as possible

Make sure that there is a continous stream of activities in the classroom

Be withit

Maintain group focus

Let students be responsible for their work

Instruction

Discuss all homework assignments

Start presentation with a review of relevant concepts

Refresh foreknowledge by means of questioning

Use during presentation concrete examples

(Table 1 continued)

Indicate (un)important passages in the textbook

Indicate what is important (oral/blackboard)

Give clear assignments for monitored practice

React positive to students

Check the assignments during monitored practice

Redirect questions and use probes

At the end of the lesson give time to students to work on their home-workassignments

Tutor at the end of the lesson students who encountered problems.

Table 2 Time engaged in the five phases* of the lesson model for the teachers in the two conditions

	Condition 1	Condition 2
Lesson phase 1	19.80	23.28
Lesson phase 2	11.72	12.35
Lesson phase 3	10.68	6.11**
Lesson phase 4	4.20	2.11**
Lesson phase 5	6.24	7.59

* In each cell of the table the mean number of intervals of 5 seconds per 5 minutes is presented.

** Difference between the two conditions is significant at a 10% level.

Table 3 The percentage of lessons with the number of lesson phases present in condition 1 and 2

	all 5 present	4	3	2	1	none
Condition 1	10.3	33.3	32.5	18.8	4.3	0.9
Condition 2	2.7	28.4	32.4	29.7	6.8	0.0